REMARKS

This paper is responsive to the Office Action mailed July 17, 2006. Reconsideration is respectfully requested. Claims 1 and 6 have been amended and claims 12 through 17 have been cancelled in order to expedite prosecution of the subject application.

1. *E-FLUENTIALS* DOES NOT TEACH USING PUBLICLY AVAILABLE DATA NOT PRESENT IN THE *E-FLUENTIALS* SURVEY, NARROWING THE PUBLICLY AVAILABLE DATA TO VARIABLES THAT MORE ACCURATELY INDICATE INDIVIDUALS OF INFLUENCE, AND APPLYING AN ALGORITHM THAT IDENTIFIES A LARGER PERCENTAGE, FROM A SMALLER POPULATION, OF INDIVIDUALS OF INFLUENCE THAN DOES *E-FLUENTIALS*

Applicant wishes to thank the Examiner for her detailed Office Action of July 12, 2006, in which every pending claim element was compared to the prior art, in this case the *e-fluentials* reference. Applicant, however, has a different interpretation of the relevance *e-fluentials* reference based on the subject invention as initially described and as now claimed.

The use of *e-fluential* type data (generally similar, but not identical, to the data referenced as "Influential" in the present application) is merely the starting point of the subject invention. This *e-fluential*/Influential type data is subject to both statistical analysis and the application of the algorithm of the subject invention in order to provide the identification of a larger percentage of "Influentials" from a smaller data pool, while using publicly available data, than does the *e-fluential*/Influential analysis. An example of this increased efficacy is provided in paragraphs 0032 and 0033 of the subject application as originally filed and published:

> [0032] For example, if a list of Influential women over the age of forty in the metropolitan New York area is desired, the following would ensue: 1) Select the .about.3,000,000 women over age forty in New York; 2) Append the Influentials bundle of predictive variables; 3) Transform (and reformat) the data into numeric representations of gains; 4) Generate

probability scores for the 3,000,000 women in the target population; and 5) Select the highest scores as those who are most likely to be Influentials and market to them.

[0033] If, for example, women within the top two scoring deciles (top 20%) were selected, this group would contain half of all the Influentials that exist in the entire population of 3,000,000 women. On the premise that 10% of the population are Influentials, we could expect that 300,000 of our New York, over-forty women are such. Targeting our top two scoring deciles would isolate half of them, or 150,000. Therefore, targeting 600,000 (20% of the 3,000,000) yields 150,000 Influentials (50% of 300,000). In the absence of the Influentials model, it would be necessary to communicate with half of all the women (1,500,000) to have contacted the same 150,000 Influentials.

The advantage of the subject invention over the *e-fluential*/Influentials methodology is also graphically shown in FIGS. 16 and 17 of the subject application as originally filed and published.

Independent claims 1 and 6, as amended, now require at the close of element e. thereof:

"wherein c. includes data of both the individuals with a greater probability of influencing others and the individuals without a greater probability of influencing others, the data of c. being merged with a plurality of publicly available data elements not present in b., the merged data being statistically processed to narrow and flag the merged data to fewer variables that indicate individuals with a greater probability of influencing others, an algorithm being derived that predicts which individuals will have a greater probability of influencing others from the publicly available data elements not present in b., and the algorithm being applied to the merged data elements to identify a larger percentage of individuals than that of d. from a smaller data population than that of d. that have a greater probability of influencing others."

Support for the above amendment to independent claims 1 and 6, as well as a better understanding of the subject invention, is best found in the paragraphs 0025 through 0031 of the application as originally filed and published:

> [0025] At step one, the data file indicating Influential/nonInfluential status is merged with approximately 900 data elements, which had been purchased from third parties in a manner generally known and available in the art and gleaned from the publicly available U.S. Census data, and each individual is matched with his or her vector of descriptive variables (the individuals surveyed to create the initial database of Influentials and nonInfluentials are, of course, identical to those individuals for whom other descriptive data has been collected by the third parties). This stage essentially involves electronically reading the merged databases, appending them to the target name and address file and converting the database into a format consistent with the analytic requirements for targeted market research. Variables are restaged in a manner that can be used for modeling (e.g., date fields are converted to numeric representations). Each positional data element is associated with a variable name.
>
> [0026] After the data are read in and transformed, the computer program randomly divides the analytic file into two components: 1) a test file; and 2) a validation file. The second, third and fourth step (infra) are conducted on the test file only, and the validation file is used for model validation in step 4.
>
> [0027] At the end of step one, the test file of Influentials and non-Influentials consists of 900 purchased, existing variables, plus the one variable indicating one's status as an Influential or nonInfluential, created by this invention, all of which are appended to each individual. Thus, at step two, the system applies a variant of chi square analysis to narrow the number of variables to the strongest (i.e., those variables where

preexisting responses to the prior survey questions most closely correlate with Influential status) fifty to seventy. A computer program automatically 1) creates intervals for continuous variables; 2) assigns response indexes to each interval for continuous, categorical and binary variables and 3) identifies variables to be retained based on their having an index greater than an established threshold and representing a percentage of the sample also greater than an established threshold. Values meeting the selection criteria are flagged and are then processed at step three.

[0028] At step three, variables are reformatted into numeric representations of gains, relative to the desired effect, so that statistical procedures can be applied without the need for onerous manual data transformations. Variables come in many forms, and may be continuous, categorical or binary. Statistical procedures require that variables are either continuous or binary. For example, a categorical variable with fifteen values must be transformed into 15 binary variables. This process is time consuming. Moreover, many continuous variables contain extreme values, which may diminish a model's predictive power and these extreme values must be smoothed. This system automatically recasts all continuous and binary data into categorical intervals and then assigns each interval a continuous, numeric value that is compatible with the requirements of the statistical procedures. The categorical data are first regrouped into "similar response" clusters and then assigned the numeric value that represents the interval's relative effect on the response of interest (i.e., whether or not someone is an Influential). These numerical values are then stored as formats that can be associated with raw (pre-transformed) data values that fall within each prescribed interval. In this way, raw data for populations to be scored for Influentials can be easily transformed to formats easily useable by the model.

[0029] For each categorical or binary value, the software creates a transformed value representing the distance between the category response percent and the sample-wide response percent. This is a linear function of how this value affects the response in question (i.e., whether or not someone is an influential). All transformed values are stored as formats that are associated with the categorical values derived above. Variables are then renamed in a manner consistent with format naming conventions.

[0030] Step four involves the application of statistical procedures to the test dataset to select the final set of predictive variables, test interaction and quadratic terms, fit the model and validate it against the random validation sample created in step two. This comprises the Influentials database scoring algorithm.

[0031] Step five involves applying the algorithm to external data for scoring. A company would seek to have its database, or a purchased database, of names and addresses processed. The processing entails appending the reduced subset of Influentials predictive variables from the third party data vendor to the database of names and addresses. The raw data values for each individual are then transformed into their corresponding values (numeric representation of gains) that were stored as formats in step 3 above. As a result of these transformations, the data have been recast in a format that maximizes predictive power and is consistent with the data format required to invoke the Influentials scoring algorithm.

Thus, the *e-fluential*/influentials methodology, being only the starting point of the subject invention, neither teaches nor suggests:

1. Using publicly available data (i.e. U.S. Census data) not present in the initial *e-fluential*/influentials survey;

2. Using statistical analysis to narrow the data of 1., above, to a group of variables associated with the individuals in the data that more accurately indicate individuals of influence; or

3. Applying an algorithm to the data of 2., above, that results in the identification of a larger percentage, in a smaller population, of people of influence than is located by just employing the *e-fluential*/influentials methodology.

2. *E-FLUENTIALS* DOES NOT TEACH APPLYING THE CREATED ALGORITHM BY AN END USER TO A SECOND, PREVIOUSLY UNUSED POPULATION OF INDIVIDUALS TO DETERMINE INDIVIDUALS OF INFLUENCE

Independent claim 6 (as amended to modify element e. thereof) at element f. thereof, now requires that the created algorithm be applied to a second, previously unused population to determine individuals of influence by an end user of the algorithm in their business:

> f. applying the assessed relationship to a second population of individuals to determine which of the individuals in the second population having a greater probability than other individuals in the second population of influencing the choices made by individuals.

The intent of element f. of claim 6 is to employ in a practical application the algorithm that was created under element e. of claim 6 as amended, as best shown in paragraph 0031 of the application as originally filed and published:

> [0031] Step five involves applying the algorithm to external data for scoring. A company would seek to have its database, or a purchased database, of names and addresses processed. The processing entails appending the reduced subset of Influentials predictive variables from the third party data vendor to the database of names and addresses. The raw

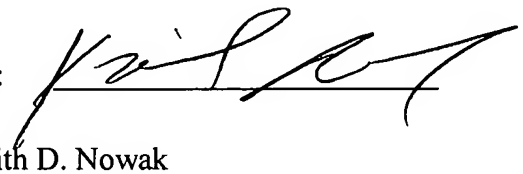
data values for each individual are then transformed into their corresponding values (numeric representation of gains) that were stored as formats in step 3 above. As a result of these transformations, the data have been recast in a format that maximizes predictive power and is consistent with the data format required to invoke the Influentials scoring algorithm.

The *e-fluential*/Influentials methodology of course does not have an end user of the algorithm in their business apply the created algorithm to a second, previously unused population to determine individuals of influence because the *e-fluential*/Influentials methodology never contemplated creating an algorithm that results in the identification of a larger percentage, in a smaller population, of people of influence than is located by just employing the *e-fluential*/influentials methodology alone.

For the above stated reasons, Applicant respectfully requests that the above rejections be withdrawn and that subject application be allowed to pass promptly to issuance.

Respectfully submitted,

Dated: January 17, 2007

By: 

Keith D. Nowak
Reg. No. 27,367
Carter Ledyard & Milburn LLP
2 Wall Street
New York, NY 10005
Telephone No.: (212) 238-8610
Facsimile No.: (212) 732-3232